On a Generic Differentiation Rule for all Mathematical Operators

Ibrahim M. Alabdulmohsin

Abstract— **In this paper, we present a generic differentiation rule that is applicable to all mathematical operators and illustrate how the generic differentiation rule is vital in deducing the derivatives of complex mathematical operations such as functional iteration. We also show how the generic rule offers more insight into the concept of differentiation and provides systematic proofs to differentiation rules that, otherwise, would have been proven using ad hoc approaches. Next, the generic rule is shown to yield interesting results including a proof that the first-degree approximation of the composite of iterated functions, at the limit where the iteration variable approaches zero, is the sum of the iterators of those functions. Consequently, at the vicinity of the specified limit, composition of iterated functions is remarkably symmetric.**

Index Terms—**Differentiation, Recursion, Iterated Functions.**

I. Introduction

A mathematical operation is an action upon which new values are computed out of one or more input values. In the special case where the number of input values is strictly two, the operation is commonly referred to as a binary operation, where examples of binary operations include addition, multiplication, and exponentiation over a given domain. The input values in a mathematical operation are sometimes called operands or arguments while outputs are frequently termed results or outcomes. Mathematical operations are not limited to numbers, however. For instance, the operands of a mathematical operation could be matrices, logical states, sets, or even entire functions.

Mathematical operators can be categorized into two distinct classes: canonical and uncanonical operators. When describing a mathematical object, the term *canonical* in mathematics generally implies that an object is in its "simplest or standard form" [1]. In this paper, canonical operators refer to mathematical operators that are fundamental; meaning that they cannot be expressed in closed-form using any combination of other canonical operators. More precisely, given a set of canonical operators $\omega=\{\omega_0, \omega_1, \ldots \omega_i\}$ and a domain D_ω that defines the allowable values of the operands of ω, an operator ω_{i+1} is canonical if and only if it cannot be expressed in closed-form using any combination of operators in ω and *constant* values in D_ω. The previous definition clearly implies that an operator is not canonical per se, but only canonical with reference to a specific domain and a given list of canonical operators. To illustrate why this is central, let us assume initially that $\omega = \phi$, which implies that any operator is initially canonical with respect to ω. If we define addition to be our first canonical operator, i.e. ω_0 is the addition operator, then the operator ω_1 defined by $A\ \omega_1\ B = AB + A$ is also a canonical operator since it cannot be expressed in closed form solely in terms of the addition operator. However, with the addition of the new canonical operator ω_1, the multiplication operator is no longer a canonical operator since $A \cdot B = (A\ \omega_1\ B)\ \omega_0\ ((-1\ \omega_1\ A)\ \omega_0\ (-1))$. Furthermore, the domain D_ω is also vital in the definition of canonical operators. Consider, for instance, the addition and subtraction operators. When D_ω is the set of all integers, the subtraction operator is uncanonical with respect to addition and multiplication because $A - B = A + (-1)\cdot B$. However, if D_ω is the set of whole numbers only, the operands of the addition operator cannot be negative by definition of D_ω and, thus, the subtraction operator is canonical. The distinction between canonical and uncanonical operators will play a key role in the subsequent sections.

In calculus, the use of differentiation rules for different mathematical operators has been adopted to avoid the, somewhat, complex computations of the derivative by definition as a limit of quotient [2]. Examples of well-known differentiation rules are the sum rule shown in (1), the product rule shown in (2), and the functional power rule shown in (3). Nonetheless, each rule is applicable to its own mathematical operator only, hence the names, which entails them insufficient in countless other scenarios. One particular example that will be examined in this paper is the continuous/partial functional iteration, which, in its utmost general form, cannot be expressed in closed-form using any combination of the three canonical operators: addition, multiplication, and exponentiation over any domain of numbers D_ω. In fact, functional iteration alone may, indeed, generate a countless number of canonical mathematical operators as will be shown next.

$$(f+g)' = f' + g' \qquad (\textit{Sum Rule}) \qquad (1)$$

$$(f\cdot g)' = f'\cdot g + f\cdot g' \qquad (\textit{Product Rule}) \qquad (2)$$

$$\left(f^{g}\right)' = f^{g}\cdot g'\cdot \ln f + f^{g-1}\cdot g\cdot f' \qquad (\textit{Functional Power Rule}) \qquad (3)$$

Functional iteration is the study of how the outcome of a function changes under repetition. In functional iteration, the output of a function is fed back as input to exactly the same function, constituting what is referred to as an *iteration*. To construct a second iteration, the new output is fed back as input to the same function again, and so on. More precisely, given a set X and let f: $X\rightarrow X$ be a function, the nth iterate of the function f, denoted as ${}^{n}f$, satisfies ${}^{n}f = f \circ {}^{n-1}f = f({}^{n-1}f)$. As discussed in details in [3], functional iteration satisfies many properties including the following:

Property 1: ${}^{a}f({}^{b}f(x)) = {}^{a+b}f(x)$

Property 2: ${}^{a}({}^{b}f(x)) = {}^{ab}f(x)$

Property 3: ${}^{0}f(x) = Id(x) = x$

Property 4: ${}^{1}f(x) = f(x)$

Manuscript received February 28, 2009.

I. M. Alabdulmohsin is a communications engineer at the Saudi Arabian Oil Company (Saudi Aramco). He is currently enrolled in the M.S. degree program in Electrical Engineering at Stanford University. **Email**: abdulmim@stanford.com

Property 5: $\;^{-1}f(x)$ is the inverse function of $f(x)$.

Iteration is a functional operator, where the operand is a univariate function *f*(*x*) and the output is its corresponding iterated function $^{n}f(x)$. Because the iterated function $^{n}f(x)$ is a function of two variables, $^{n}f(x)$ can be modeled as a binary operator ω_f over a domain of numbers D_ω such that $^{n}f(x) = n\ \omega_f\ x$. With this in mind, functional iteration could potentially yield an infinite number of canonical operators. For instance, the iterated trigonometric functions $^{n}\sin(x)$ and $^{n}\cos(x)$ are expressions of infinite sums and cannot be expressed in closed-form using addition, multiplication, or exponentiation, lending them canonical with the respect to those operators.

In this paper, we adopt an abstract view of mathematical operators and prove a generic differentiation rule that is applicable to all mathematical operators. More specifically, we address the following question: given the derivatives $\frac{\delta}{\delta x}\left(A\ \omega\ f(x)\right)$ and $\frac{\delta}{\delta x}\left(f(x)\ \omega A\ \right)$, where *A* is a constant, and ω is a generic operator that is not necessarily symmetric, how do we compute the derivative $\frac{\delta}{\delta x}\left(g(x)\ \omega\ f(x)\right)$? If the number of binary canonical operators were finite, an answer to such question would not be needed if a complete set of differentiation rules for all binary canonical operators could be deduced. However, as discussed earlier, the number of binary canonical operators is, most probably, infinite. The paper starts with an answer to the question above and shows how the three differentiation rules, stated earlier, can be derived easily using a systematic approach. Next, the generic differentiation rule is used to deduce interesting results regarding functional iteration.

II. A Generic Differentiation Rule

Theorem 2.1: Given a generic mathematical operator ω, $\frac{\delta}{\delta x}\left(g(x)\ \omega\ f(x)\right) = \frac{\delta}{\delta x}\left(g\ \ \omega\ f(x)\right) + \frac{\delta}{\delta x}\left(g(x)\ \omega f\ \right)$, where *g*(*x*) and *f*(*x*) are assumed variable during differentiation and *g* and *f* are assumed constant.

Proof:

By definition of differentiation:

$$\frac{\delta}{\delta x}\left(g(x)\ \omega\ f(x)\right) = \lim_{h\to 0}\frac{g(x+h)\ \omega\ f(x+h) - g(x)\ \omega\ f(x)}{h}$$

$$= \lim_{h\to 0}\frac{g(x+h)\ \omega\ f(x+h) - g(x+h)\ \omega\ f(x) + g(x+h)\ \omega\ f(x) - g(x)\ \omega\ f(x)}{h}$$

$$= \lim_{h\to 0}\frac{g(x+h)\ \omega\ f(x+h) - g(x+h)\ \omega\ f(x)}{h} + \lim_{h\to 0}\frac{g(x+h)\ \omega\ f(x) - g(x)\ \omega\ f(x)}{h}$$

$$= \lim_{h\to 0}\frac{g(x)\ \omega\ f(x+h) - g(x)\ \omega\ f(x)}{h} + \lim_{h\to 0}\frac{g(x+h)\ \omega\ f(x) - g(x)\ \omega\ f(x)}{h}$$

$$= \frac{\delta}{\delta x}\left(g\ \ \omega\ f(x)\right) + \frac{\delta}{\delta x}\left(g(x)\ \omega\ f\ \right)$$

The generic differentiation rule in theorem 2.1 is a very insightful tool. It implies that the derivative of a function *f*(*x*), where the independent variable *x* appears more than once, can be computed by partitioning all appearances of the independent variable into two groups such that one group is assumed constant during differentiation in one time and the other group is assumed constant during differentiation in a second time. The overall derivative of the function *f*(*x*) is, then, the sum of the two results.

The generic differentiation rule can be used to prove the differentiation rules stated earlier in this paper. To prove the sum rule shown in (1):

$$= \frac{\delta}{\delta x}\left(f(x) + g(x)\right) = \frac{\delta}{\delta x}\left(f(x) + g\right) + \frac{\delta}{\delta x}\left(f + g(x)\right)$$

However, the derivative of *f*(*x*)+*A*, where *A* is constant, is equal to the derivative of *f*(*x*), which can be proven directly using the definition of derivative as a limit of quotient. Thus,

$$= \frac{\delta}{\delta x}\left(f(x) + g(x)\right) = \frac{\delta}{\delta x} f(x) + \frac{\delta}{\delta x} g(x)$$

For the product rule, we know that $= \frac{\delta}{\delta x} af(x) = a\frac{\delta}{\delta x} f(x)$, which can be proven directly by definition of differentiation. Therefore, using the generic rule:

$$= \frac{\delta}{\delta x}\left(f(x)g(x)\right) = \frac{\delta}{\delta x}(g \cdot f(x)) + \frac{\delta}{\delta x}(f \cdot g(x))$$

$$= g\frac{\delta}{\delta x} f(x) + f\frac{\delta}{\delta x} g(x)$$

The functional power rule is even more interesting. Typically, such rule is proven using properties of logarithms. A typical proof would go as follows:

$$y = f(x)^{g(x)}$$

$$\ln(y) = g(x)\ln f(x) \Rightarrow y^{-1}\frac{\delta}{\delta x} y = g'(x)\ln f(x) + \frac{g(x)}{f(x)} f'(x)$$

$$\frac{\delta}{\delta x} y = f(x)^{g(x)}\left(g'(x)\ln f(x) + \frac{g(x)}{f(x)} f'(x)\right)$$

However, using the generic rule, no prior knowledge of the properties of logarithms is needed. At first, assume *g*(*x*) is constant. Thus,

$$= \frac{\delta}{\delta x} f^{g}(x) = g \cdot f^{g-1}(x) \cdot f'(x)$$

In the second time, we assume *f*(*x*) to be constant. Thus,

$$= \frac{\delta}{\delta x} f^{g(x)} = f^{g(x)}(\ln f)(g'(x)$$

Consequently, the overall derivative is the sum of the previous two results:

$$\frac{\delta}{\delta x} f(x)^{g(x)} = g(x) \cdot f^{g(x)-1}(x) \cdot f'(x) + g(x) \cdot f^{g(x)-1}(x) \cdot g'(x)$$

, which is equivalent to the well-known functional power rule. Note that in all three cases, the proofs followed the same systematic approach, namely to hold one function constant during differentiation in one time, hold the other function constant during differentiation in a second time, and, finally, sum up the two results. Alternatively, proofs to the differentiation rules would have been accomplished using ad hoc approaches. It is, therefore, obvious that such a systematic approach provides more insight into the concept of differentiation. In addition, the generic rule is imperative when the function involves canonical operators other than ones discussed so far such as functional iteration.

III. Differentiation and Iteration

In [3], the author has shown that a function *g*(*n*, *x*) is an iterated function $^{n}f(x)$ if and only if it satisfies the following two conditions:

1. $g(0, x) = x$

2. $\frac{\delta}{\delta n} g(n,x) = [\lim_{n\to 0} \frac{\delta}{\delta n} g(n,x)] \cdot \frac{\delta}{\delta x} g(n,x)$

The quantity $\lim_{n\to 0} \frac{\delta}{\delta n} g(n,x)$, called the iterator of $f(x)$ and denoted $h_f(x)$, is a unique property of $f(x)$ and satisfies many interesting properties including: $\frac{d}{dn} {}^n f(x) = h_f({}^n f(x))$. With this in mind, the relation between the derivative $\frac{\delta}{\delta x} {}^x f(x)$ and the iterator $h_f(x)$ can be determined using the generic differentiation rule.

Theorem 3.1: $\frac{\delta}{\delta x} {}^x f(x) = h_f({}^x f(x)) \left(1 + \frac{1}{h(x)}\right)$, where $h_f(x)$ is the iterator of $f(x)$.

Proof: The function ${}^x f(x)$ has two appearances of the independent variable x. Thus, the derivative $\frac{\delta}{\delta x} {}^x f(x)$ can be determined by holding one appearance of the independent variable constant at a time and summing up the two results. Holding g=x constant, $\frac{\delta}{\delta x} {}^x f(g) = h_f({}^x f(g)) = h_f({}^x f(x))$.
Holding the second appearance of the independent variable constant and knowing that $\frac{\delta}{\delta n} {}^n f(x) = h_f(x) \cdot \frac{\delta}{\delta x} {}^n f(x)$,

$$\frac{\delta}{\delta x} {}^g f(x) = \frac{h_f({}^g f(x))}{h_f(x)} = \frac{h_f({}^x f(x))}{h_f(x)}.$$

Thus,

$$\frac{\delta}{\delta x} {}^x f(x) = h_f({}^x f(x)) \left(1 + \frac{1}{h_f(x)}\right)$$

Theorem 3.1 can be tested using well-known closed-form expressions of iterated functions as illustrated in the following two examples. In both examples, it is clear that theorem 3.1 holds true as expected.

Example 3.1: If $f(x) = x + k$, then

$${}^n f(x) = x + kn \Rightarrow {}^x f(x) = (1+k)x \Rightarrow \frac{\delta}{\delta x} {}^x f(x) = (1+k).$$

Using theorem 3.1,

$$h_f(x) = k \Rightarrow h({}^x f(x)) \left(1 + \frac{1}{h(x)}\right) = k(1 + 1/k) = (1+k) = \frac{\delta}{\delta x} {}^x f(x)$$

Example 3.2: If $f(x) = kx$, then

$${}^n f(x) = k^n x \Rightarrow {}^x f(x) = k^x x \Rightarrow \frac{\delta}{\delta x} {}^x f(x) = k^x (1 + (\ln k) x)$$

Using theorem 3.1,

$$h_f(x) = (\ln k) x \Rightarrow h({}^x f(x)) \left(1 + \frac{1}{h(x)}\right) = k^x (1 + (\ln k) x) = \frac{\delta}{\delta x} {}^x f(x)$$

In the next theorem, the importance of the generic differentiation rule is demonstrated, again, in computing the composition of iterated functions at the limit where the iteration variable approaches zero.

Theorem 3.2:

$$\lim_{n\to 0} {}^n f_1 \circ {}^n f_2 \circ {}^n f_3 \circ \ldots \circ {}^n f_k(x) = x + \left(\sum_{j=1}^{k} h_{f_j}(x)\right) n$$, where

$h_{f_j}(x)$ is the iterator of $f_i(x)$.

Proof: Defining g=n and holding g constant, and using the definition of iterators:

$$\lim_{n\to 0} \frac{\delta}{\delta n} {}^n f_1 \circ {}^g f_2 \circ {}^g f_3 \circ \ldots \circ {}^g f_k(x) = h_{f_1} \circ {}^g f_2 \circ {}^g f_3 \circ \ldots \circ {}^g f_k(x)$$

The constant g, however, is held constant at the limit g→0, while implies:

$$\lim_{n\to 0} \frac{\delta}{\delta n} {}^n f_1 \circ {}^g f_2 \circ {}^g f_3 \circ \ldots \circ {}^g f_k(x) = h_{f_1} \circ {}^0 f_2 \circ {}^0 f_3 \circ \ldots \circ {}^0 f_k(x) = h_{f_1}(x)$$

Differentiating with respect to the other partition of n,

$$\lim_{n\to 0} \frac{\delta}{\delta n} {}^g f_1 \circ {}^n f_2 \circ {}^n f_3 \circ \ldots \circ {}^n f_k(x) = \lim_{n\to 0} \frac{\delta}{\delta n} {}^0 f_1 \circ {}^n f_2 \circ {}^n f_3 \circ \ldots \circ {}^n f_k(x)$$

$$= \lim_{n\to 0} \frac{\delta}{\delta n} {}^n f_2 \circ {}^n f_3 \circ \ldots \circ {}^n f_k(x)$$

Thus, the overall derivative at the specified limit is given by:

$$\lim_{n\to 0} \frac{\delta}{\delta n} {}^n f_1 \circ {}^n f_2 \circ {}^n f_3 \circ \ldots \circ {}^n f_k(x) = h_{f_1}(x) + \lim_{n\to 0} \frac{\delta}{\delta n} {}^n f_2 \circ {}^n f_3 \circ \ldots \circ {}^n f_k(x)$$

The same process can be repeated with the second quantity $\lim_{n\to 0} \frac{\delta}{\delta n} {}^n f_2 \circ {}^n f_3 \circ \ldots \circ {}^n f_k(x)$ to arrive at:

$$\lim_{n\to 0} \frac{\delta}{\delta n} ({}^n f_1 \circ {}^n f_2 \circ {}^n f_3 \circ \ldots \circ {}^n f_k(x)) = \left(\sum_{j=1}^{k} h_{f_j}(x)\right)$$

Since $\lim_{n\to n_0} {}^n f(x) = {}^{n_0} f(x) + \left(\lim_{n\to n_0} \frac{\delta}{\delta n} {}^n f(x)\right) n$ and knowing that ${}^0 f_i(x) = x$:

$$\lim_{n\to 0} {}^n f_1 \circ {}^n f_2 \circ {}^n f_3 \circ \ldots \circ {}^n f_k(x) = x + \left(\sum_{j=1}^{k} h_{f_j}(x)\right) n$$

As a result, at the vicinity around the limit where the iteration variable approaches zero, composition of iterated functions becomes surprisingly symmetric. Theorem 3.2 can be tested using well-known closed-form expressions of iterated functions as the following example illustrates.

Example 3.3: Suppose $f_1(x) = x + k$ and $f_2(x) = px$, where both k and p are constants, then

$${}^n f_1 \circ {}^n f_2(x) = p^n x + kn.$$

Thus,

$$\lim_{n\to 0} \frac{\delta}{\delta n} ({}^n f_1 \circ {}^n f_2(x)) = (\ln p) x + k$$

, which implies that:

$$\lim_{n\to 0} {}^n f_1 \circ {}^n f_2(x) = x + ((\ln p) x + k) n.$$

On the other hand,

$${}^n f_2 \circ {}^n f_1(x) = p^n (x + kn).$$

Thus,

$$\lim_{n\to 0} \frac{\delta}{\delta n} ({}^n f_2 \circ {}^n f_1(x)) = (\ln p) x + k$$

, which, again, implies that

$$\lim_{n\to 0} {}^{n}f_2 \circ {}^{n}f_1(x) = x + \big((\ln p)x + k\big)n \,.$$

In the previous example, both results are equivalent, which matches the conclusions proved earlier in theorem 3.2, and similar results are also obtained when using other combinations of iterated functions. It is worth noting that when $f_2(x)$ is the inverse function of $f_1(x)$, the iterators $h_{f_1}(x)$ and $h_{f_2}(x)$ are related to each other by $h_{f_1}(x) = -h_{f_2}(x)$ as shown in [3]. Thus, the sum of the iterators is always zero, which is consistent with the fact that the composition of their iterated functions is always the identity function $Id(x)$.

IV. Conclusion

Because the number of canonical or fundamental mathematical operators is likely to be infinite, generic mathematical rules that are applicable to all mathematical operators are of paramount importance. One particular example is the generic differentiation rule presented in this paper, which demonstrates its usefulness in providing systematic proofs to well-known differentiation rules that would have been proven using ad hoc approaches otherwise, adding further insight into the concept of differentiation. The generic differentiation rule is also helpful in deducing interesting results regarding complex mathematical operators such as functional iteration. One result, of particular interest, that can be proven using the generic differentiation rule is that the composition of iterated functions becomes symmetric at the limit where the iteration variable approaches zero.

References

[1] Dictionary.com, "canonical," in *Dictionary.com Unabridged (v 1.1)*. Source location: Random House, Inc. http://dictionary.reference.com/browse/canonical. Available: http://dictionary.reference.com. Accessed: February 20, 2009.

[2] G.B. Thomas and R.L. Finney, *Calculus and Analytic Geometry*, 9th ed. Addison Wesley Pub Co Inc, 1996.

[3] Alabdulmohsin, Ibrahim M. “Theorems and methods on partial functional iteration”, proceedings of *ICAEM09 - World Congress on Engineering*, London (2009).